DOCUMENT RESUME

ED 244 723 PS 014 281

AUTHOR House, Ernest R.
TITLE Scientific and Humanistic Evaluations of Follow Through.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Jan 81
NOTE 27p.; Paper presented at the Learning Research and Development Center/National Institute of Education Conference on Follow Through (Pittsburgh, PA, March 12-13, 1980).
PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Early Childhood Education; *Evaluation Methods; *Humanities; Program Development; *Program Evaluation; *Research Problems; *Scientific Methodology; Teacher Attitudes
IDENTIFIERS Bank Street College of Education NY; *Phenomenological Evaluation; *Project Follow Through

ABSTRACT

The thesis of this paper is that the humanistic mode of inquiry is underemployed in evaluation studies and the future evaluation of Follow Through could profitably use humanistic approaches. The original Follow Through evaluation was based on the assumption that the world consists of a single system explainable by appropriate methods; the evaluation presumed that in the domain of early childhood education there exists an underlying internally consistent reality to be discovered. The evaluation foundered badly. An alternate method of evaluation is to deal with the world in terms of ordinary language and as it appears phenomenologically to individuals. This mode of investigation emphasizes experience as lived. Ways of investigating such experiences in a disciplined manner exist; these disciplines are called the humanities. Whereas science looks for causes, the humanities look for reasons. The essence of humanistic thought seems to be that human action is intelligible only when we see it through the eyes of the acting agent, only when we see why he would do what he did. Many of the features of humanistic inquiry can be seen in Zimiles and Mayer's (1980) study of the Bank Street Follow Through Project. (RH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED244723

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

PS

Scientific and Humanistic Evaluations of Follow Through

Ernest R. House

CIRCE
University of Illinois

January 1981

This paper was prepared for the LRDC-NIE Conference on Follow Through in Pittsburgh, March 12-13, 1980.

PS 014281

Scientific and Humanistic Evaluations of Follow Through

by Ernest R. House, CIRCE, U. of Ill.

I.

When my children were a bit younger, both went to Yankee Ridge Elementary School in Urbana. A few years prior to Kristin, my oldest one, entering kindergarten there, Yankee Ridge had been a trial site for the teaching methods and materials that Carl Bereiter and Siegy Engelman had been developing, the approach later known as Direct Instruction and commercially marketed as Distar.

The kindergarten teacher had been involved in the early try-outs. She was a vivacious, energetic woman from Texas, and she was very enthusiastic about the choral reading and dramatic aspects of the Distar program. It was, she claimed, superior to any other reading program that she had encountered. She prided herself on running the only "academic" kindergarten in town and had inspired some of the other primary teachers to invest themselves personally in the Direct Instruction approach. There was a small devoted cadre of perhaps three teachers at that time.

My daughter responded well to this approach. She had always been small, shy, and quiet, a spectator of events rather than a participant. In nursery school she had most enjoyed the singing of songs and the nursery school teacher playing the guitar. She readily took to the group chanting of Direct Instruction. The dynamism of her teacher held her attention somewhat. There was a playful quality to some of the lessons. I would not say she was enthusiastic--she was never enthusiastic about school--but it was not bad.

I might also note that the Direct Instruction program, which was developed with a group of special education students, was there employed in an upper-middle class school with a mean I.Q. above 125 on the California. In Follow Through it was used with disadvantaged students and was advertised by its commercial publishers as a panacea for all students with learning problems. Opportunity knocks several times apparently.

Another part of the Yankee Ridge primary program was math, in which the teachers used IPI materials. Several years before, when I was a consultant with the Illinois Gifted Program, I had been partially responsible for introducing IPI materials into the Urbana Schools. In a search for new programs to exhibit in the Illinois demonstration centers, a team of us had visited Oak Leaf School before it was officially open to visitors. On return we prompted the Illinois Department of Education to grant $10,000 in seed money to five Illinois school districts interested in adopting the Pittsburg materials. The Urbana district was one of these, and several years later the IPI materials were still in use at Yankee Ridge. Little did I know then that my own as-yet-unborn children would be instructed by a program I had helped import. Perhaps all educational reformers should be subject to such a discipline.

From the beginning my daughter was not keen about the IPI materials. They consisted mainly of worksheets so that the student could work at her own pace, a feature which presumably motivated the student by providing successful experiences, a little "reinforcement" as they say. However, what my daughter saw was that as soon as she did one sheet, she had to do another one. This was not positive motivation for her. She resisted, lagged, and complained bitterly about the materials.

This program, I might add, was developed originally in a working class school, transferred to such places as this demonstration center for gifted youth, and later used as a basic model for educating disadvantaged youths in the Follow Through experiment. (Our curriculum development efforts seem to have great versatility, whatever else they may possess.)

Four years later my son entered Yankee Ridge, a different personality altogether--aggressive, highly motivated, intensely competitive. The previous semester we had spent in England, and there he had entered a classroom in which all the other children had been reading for some time. That classroom was very old-fashioned, taught the way I had been taught 30 years ago with "Dick and Jane" readers.

Not wanting to be behind, my son resolutely set out to learn to read on his own, doggedly pursuing the teacher after school. (She had to see that he had mastered each little reader before he could advance to the next one, and she never had enough time during class hours to attend to him.) Within three months he learned to read and closed the gap between himself and the rest of the class.

"Too bad the girl hasn't had the same preparation the boy has had," said the bald English headmaster, who also made the tapioca pudding for the students' lunch everyday. In fact, both children had gone to the same nursery school, and both had had the same teachers there.

Back in Yankee Ridge in kindergarten and in the primary grades, Colby was well ahead of most of the other children in both reading and math. The slow pace of the choral reading bothered him. It was too slow. The dynamic teacher who had been the champion of the Distar materials had

gone off to seek national office in the teacher organization. The remaining teachers followed the Distar prescriptions faithfully, but they were neither inspired nor inspiring. Something was missing. The lessons took on a routine, repetitious character for my son.

In math the better students used the linearly-sequenced IPI materials like a race track. Every one knew exactly where everyone else was on the track. The point was to race through the materials as quickly as possible. The materials had been introduced several years before, and since then the school district had fallen upon much harder financial times. There were not sufficient clerks to score the tests, which were essential to progress in the curriculum. The teachers were becoming increasingly disenchanted because of the extra record-keeping load and their inability to keep up. The children were frustrated by having to wait so long for someone to examine their lessons so they could advance. They sat for long periods of time with their hands raised, trying to get someone to look at their papers. As a parent, of course, I was not entirely happy with any of these instructional arrangements for either daughter or son. For the middle-class parent no education can be good enough for one's own child, short of the teacher devoting herself exclusively to that child. The middle-class parent seeks not learning but advantage.

What do these personal reminiscences have to do with evaluating Follow Through? It seems to me that a proper evaluation should capture some of the complexities of such a real, live situation. Certainly, the individual differences among children is obvious. The different social contexts in which the programs operate are less so. That these programs will have subtle and profound effects not recorded by traditional achievement measures is highly

probable. In implementation it is likely that schools will adopt part of a program, or in this instance, parts of two different programs. Furthermore, the programs will vary significantly in their implementation, depending upon the teachers themselves. Not even a program with a written script like Distar can escape variation. Finally, programs developed in one type of setting will be used in entirely different settings with quite different students, adding a further dimension of uncertainty to the implementation and the effects, if indeed such an entity as "a program" even exists. How do evaluations account for these complexities, and how do they inform us in such a way as to be meaningful?

II.

How was the original Follow Through evaluation actually conducted? Thirteen models of early childhood education, including Direct Instruction and IPI, were compared to one another. Each sponsor or model was assigned several sites, i.e. school districts in which to implement its program. Within each site Follow Through and corresponding control classes were chosen at the end of 3rd grade. Those children who were still left in the classes were administred four psychometric measures--the 3rd grade Metropolitan Achievement Test, the Intellectual Achievement Responsibility Scale, the Coopersmith Self-Esteem Inventory, and the Raven's Coloured Progressive Matrices Test. Using previously collected data such as entry achievement scores as covariates, the Follow Through classes were compared to their controls to see which had statistically significant higher scores on the outcome measures. Findings for individual classes and sites were added together in a complicated way to provide a summary score for each Follow

Through model. These summary scores could then be compared to those of other models to determine the winners and losers.

As is well known, the publication of the evaluation findings was greeted by a barrage of hostile criticism, most of which attacked the validity of the findings in some way. There were far more attacks than defenses of the study, and I think it accurate to say that the study is widely believed to be a bad example of what an evaluation is supposed to be. There is less agreement as to why the evaluation went wrong. Some critics point to technical deficiencies, others to political and historical factors. To this list of potential sources of error, I would like to add a more esoteric one: the particular notion of science upon which the evaluation was premised. Much of the evaluative reasoning was based upon the techniques and methods of mainstream social science, the aspirations of which have been to emulate the methods of the physical sciences.

Our notion of what science should be has evolved historically from the Enlightenment, from the birth of modern science in the 17th and 18th centuries. In their reaction against medieval scholasticism, which tried to discover the place of all things within God's purpose, the scientists and philosophers of the Enlightenment conceived the universe as a single, undifferentiated substance. Apparent changes within this substance, changes one could perceive with the senses, were actually changes in state, and these changes could be modeled by mathematical formulas. Newton's physics was the paradigmatic example for later generations.

Among visible phenomena, one could discover underlying and more fundamental relationships, and these discoveries would culminate in a coherent system of elements. Thus, underlying external appearances and differences,

there was a logically-connected structure of physical laws, and it was only a matter of time until scientists discovered what these laws were. Although Renaissance scientists had given up the notion of Divine Will ruling the universe, they had not given up the idea that there was an underlying order within it.

Given this conception of the nature of reality, it followed (particularly for those like Leonardo, Copernicus, and Galileo, who had rediscovered Plato and Pythagoras [Abbagnano, 1967]) that scientific explanation should be abstract, preferrably mathematical. One could apply mathematical techniques to that which was measurable, and that which was measurable was a subset of what was observable. Application of the appropriate method would lead to the formulation of general laws. As mentioned, Newton was the supreme example of the scientist in action--observation leading to the mathematical formulation of universal laws of reality.

Success with these methods was rapid, and within the next few centuries, the confidence of science steadily increased. Observation and experimentation were advanced by some as the sole reliable method of knowledge. Exact measurement was important because some scientific propositions could be tested only by careful calculation. Science was conceived as a single coherent body of logical conclusions arrived at by universally valid principles of thought, such as deduction. These conclusions were founded securely upon controlled observation and experiment.

Ordinary language was viewed with suspicion. The goal of science, escaping from the clutches of religion, was to eliminate superstition, bias, and emotion from observations of the natural world. Ordinary language was laden with such baggage. Hence, scientific language should be shorn of biases and emotion. When possible, it should even be quasi-mathematical.

Isaiah Berlin (1980) has formulated three assumptions upon which modern science rests. First, science assumes that every question has one and only one true answer. If one does not arrive at such an answer, then one has asked the wrong question. Asking the proper one will yield the right answer. Second, there is one method or set of methods for discovering the answer, and this method is rational in character. Often the method is construed as observation and experiment or as the hypothetical-deductive process. Furthermore, this method for discovering truth is identical across all fields, although it may differ in detail somewhat. Third, the answers discovered by such a method are true universally, true for all people in all times and all places. Truth is not relative in any way. Underlying all three assumptions is the notion that the world consists of a single system explainable by the appropriate methods.

How do these ideas derived from physical science apply to the study of humans and human society? Within this Enlightenment tradition humans are seen as objects in nature. Human nature is regarded as being essentially similar in all times and places. Even though circumstances may change in detail, there are universal human goals and patterns, and these patterns are discoverable by proper scientific methods. Local and historical variations are relatively unimportant. In other words, human reality can be studied like physical reality.

Many of these features are seen in the science of linguistics, e.g. in Chomsky's search for a universal grammar. In spite of obvious differences in human languages, Chomsky believes that underlying these differences is a single specific universal grammar which enables all humans to speak and also limits how they do so (Chomsky, 1977). Furthermore, Chomsky believes

that this universal grammar is biologically based, that is, grounded in physical principles. So even the most human of all characteristics--language--is subject to scientific investigation.

What does all this have to do with the Follow Through evaluation? The Follow Through evaluation was based upon the methods of mainstream social science, and through that, upon the notions of Enlightenment science, for mainstream social science has tried to emulate the methods of the physical sciences. The Follow Through evaluation shared similar presumptions.

Most fundamentally it was presumed that there was an underlying internally-consistent reality to be discovered. There were different approaches to early childhood education, and these approaches could be tested against one another to see "which worked best" for disadvantaged children. There was an underlying set of relationships among the models to be discovered, and once they were discovered policy makers would know how to act. After all, what would be the point of disobeying natural laws, as in physics for example?

In order to discover these relationships, one had to conduct an experiment, a very large experiment. Careful calculation, measurement, and observation were called for. Only by the appropriate methods could the answer or solution be truly determined. Great attention was paid by the evaluators to the precision of their methods. When the relationships were discovered, they could be expressed in mathematical terms. Findings about each class, site, and model would culminate in a single, coherent set of findings to yield a clear picture of which model was best. The whole thing would add up.

Furthermore, the findings would be true universally, at least for disadvantaged children across the country. If Direct Instruction or IPI were the best models, they would be best in all sites and in other cities as yet untried, just as physical laws are true. They would also be true next year and last year, from cohort to cohort. In other words, local and historical variations from site to site or from time to time were relatively unimportant. A model would achieve similar effects in similar settings. Finally, the evaluation findings should be expressed in neutral, even quasi-mathematical language, so the results would be unbiased, unemotional, and clear to everyone.

In terms of Berlin's three assumptions, there was one answer to be discovered, "Which model works best?" When the evaluators waivered on this, they were aggressively sent back to the one question. There was one method for arriving at the answer. Pleas for other methods of investigation were resolutely rejected by the evaluators and government planners. The findings were taken to be universally true, good for all times and places.

Of course, the evaluation foundered badly. The findings were unclear. Local and historical variations proved to be quite important in both implementation and results. Findings were inconsistent from one site to another. Models did well in one setting but not in others. Results even changed from one year to the next with the same model on the same site. The methods did not yield clear, coherent, and consistent results. In fact, one could obtain dramatically different results by using slightly different methods. Even choosing the site, class, or student as the unit of analysis resulted in significant differences. And the evaluation report itself, although largely statistical in content, capitalized upon emotional terms like "basic skills" to inscribe itself on the public consciousness.

Most critics said that the evaluators had chosen the wrong methods or that they had implemented the correct methods improperly or that they had generalized their results too far beyond particular sites. All of these may be true, but there is a more fundamental problem with the evaluation: The basic presumptions upon which the study was based proved to be incorrect. There is a sense in which social reality differs substantially from physical reality.

The nature of these differences I leave untouched here other than to suggest that if one evaluates an automobile for gas mileage, one will expect the gas mileage to be the same in other parts of the country (though even here there are conflicting methods of estimation). But if one institutes a new school program, one should not expect the results to be the same in other parts of the country.

III.

What does this mean? Should we abandon attempts to evaluate Follow Through altogether? It seems to me there are other possibilities. The simplest is to constrain the reach of our scientific methods in various ways. We might apply these methods where we think they have good chance of success. For example, there is no reason "a priori" to believe that Chomsky may not be correct about a universal grammar. It remains an unproved but empirical question. In scientific evaluation the closer one comes to physical phenomena, the more likely the success of these scientific methods. The less physical and the more cultural the entity being evaluated, the less likely the success with these methods. In other words, we might apply scientific methods more wisely, realizing that they do not

generate the total truth.

Rather than pursue that idea, in the rest of this paper I would like to explore a mode of inquiry other than the scientific. What if one does not presume that there is one true method or that results are universal or even that there is only one true answer? Another way of investigating (and evaluating) is via the "manifest image" of humans, that is, by dealing with the world in terms of ordinary language, dealing with the world as it appears phenomenologically to individuals.

For example, evaluations might be case studies in which the evaluator constructs a narrative, a story as it were, as to what the program is all about. The story may be told in the words of the participants of the program. A major presumption of this approach is that one has to know what has happened and is happening within the program to know what is possible for its future. Every program, like every person, is bound to its past. To understand a program fully, one must see it through the eyes of its participants because it has a significance and meaning to them that only they understand. Even though the evaluator or investigator may not agree with the meaning which participants ascribe to their actions, the evaluator may believe it is vital to know what these meanings are.

One problem with the scientific mode of investigation is that it does not capture in any recognizable way the experiences of the participants. Hence, one cannot fully appreciate the actions they might take. These experiences are recorded in ordinary, everyday language and concepts. If, on the one hand, the scientific mode of investigation emphasizes method as a way of controlling biases (a reasonable expectation derived from the historical conditions preceding the Enlightenment)

this mode of investigation emphasizes experience as lived. For example, in the stories and anecdotes that I told about my children at the beginning of this paper, one can discern bases for action that children, teachers, parents, and even program developers might take as a result of these experiences. There is virtually nothing in the original scientific evaluation of Follow Through that would give a hint as to the occurence of such events. Science does not record experience in such ordinary terms. Yet it is undeniable that the nature of this experience determines action in a substantial way.

If the scientific findings of the original Follow Through evaluation really did reflect an underlying immutable, universal reality in the way that laws of physics do, then one might draw some basis for action from them. There is no sense in disobeying the laws of physics. But it is precisely this presumption of similarity between physical and social reality that I am denying. The findings of the Follow Through evaluation do not have the necessity of the findings of physics even though elaborate quantitative methods were employed, and this is not simply because the wrong methods were employed.

What about the anecdotes about my children? Surely one cannot accept them at face value. They are undocumented and unsubstantiated. They may well be wrong. In fact, in the sense that they were conceived by a concerned parent, there is the possibility that they are distorted to fit the self-interests of the parent and the children. In accepting the ordinary language and experiences of participants, one has reintroduced the possibilities of bias that scientific method was designed to exclude. If the

strength of this experience-based mode of investigation is that it captures the valued life experiences of participants, and hence provides an exceptional guide to their actions, its weakness is that it also mixes in their biases and those of the investigator. It is sometimes biased, subjective, and undisciplined.

There are ways of investigating such experiences in a disciplined manner, however, and these disciplines are called the humanities. The humanities share a concern for the "manifest image" of humans, the analysis of everyday experience in terms of individual human beings and their actions. These actions are represented verbally, including the utterances of the individuals themselves. The humanities as fields of study also grew out of the Renaissance. Humanists also struggled for freedom from the hierarchy of the church and feudalism. This they hoped to achieve by resurrecting the study of the ancient manuscripts of Rome and Greece. This effort led to historical and textual studies. In a sense, the humanists tried to recapture the mental powers that they thought the ancients had acquired from poetry, rhetoric, history, ethics, and politics (Abbagnano, 1967).

The aim of the humanists was to assert the value and dignity of humans, to reintegrate humans back into the scheme of nature and history. "Man is the measure of all things," was their adopted motto. Aristotelianism, particularly the physics and logic, had long served as a cornerstone of medieval society, but the humanists preferred Aristotle's moral philosophy over his physics. Although both scientists and humanists started from the same origins, within a few centuries their paths diverged.

Against the extreme rationalism of scientific method, particularly as expounded by Descartes, who debunked studies like history altogether, came the Counter-Enlightenment. The main figure was Vico who first denied some of the scientific presumptions as they applied to the social and cultural world. To study humans as natural objects in the scientific mode did not make sense, said Vico (Berlin, 1980). Human activities were totally knowable from the "inside", as it were. Being human, one had insight into other people's motivations and purposes, thereby permitting a deeper and superior knowledge. There was no need to study humans as objects.

Furthermore, according to Vico, there was no unchanging human nature as presumed by the scientists. Humans changed with their culture, and it was no good reading current culture into that of the past. Cultures change, and the only way to understand the past was by an examination of past language, myths, and rites. In other words, one had to treat past societies as different in important ways. Scientific findings about humans were not necessarily true for all times and places. Nor was there only one true method for achieving such insights. Scientific methods were fine when applied to the natural world, but not when applied to humans.

Vico's analysis eventually led to theories of historical evolution and cultural relativism. The core idea was the notion of what a culture was, of it's unity and uniqueness (Berlin, 1980). To understand a culture one must possess imaginative insight and be able to conceive of more than one way of categorizing reality. History should be an account of the variety of experience. Only by tracing the genesis and history of a culture could it be properly understood. In a sense there was not one great

underlying coherent reality to be discovered but many. Knowledge was not cumulative in the scientific sense of progressively discovering the one underlying reality with one method. Also in contrast to scientific method, informed imaginative insight was necessary to understanding. One gained knowledge by experience or by imagination. There was no single structure of reality and no timeless "natural law."

As it has come down to us, the scientific mode of inquiry presumes to give us the world as an entity independent of our way of perceiving it, in other words, objectively. Since one can cumulatively uncover the single underlying reality, scientific knowledge will be progressive (Scruton, 1980). The humanistic mode of inquiry, on the other hand, presumes to give us the world as it is experienced by other people, as seen through their cultural lenses. The essence of humanistic thought seems to be that human action is intelligible only when we see it through eyes of the agent who does it, only when we see why he would do what he did. Only then do we see the order in the other person's experience. Science looks for causes as explanations; the humanities look for reasons. Meaning and significance are derived from the human context within which the action occurs. The significance of events is a "felt" significance. It is not independent of our ways of perceiving it. In fact, it is very much part of our way of perception and, is in that sense, subjective.

The humanities, however, are simply collections of feelings. They presume to be disciplined inquiry. In this there is agreement that they involve higher conceptual powers of humans, particularly the capacity to make comparisons and impose order on experience (Scruton, 1980), and the capacity for making distinctions of value, including moral distinctions (Olafson, 1979). Value, order, and coherence are terms one hears

often from humanists. It is not difficult to see that the fundamental concepts of the humanities might be useful to formal evaluation. In fact, I would say that the humanistic mode of inquiry is one way of doing evaluation, and one that differs substantially from the scientific mode.

Besides seeing humans as the possessor of certain mental powers, the humanities portray individual persons and their actions in terms of their intentions and purposes (Olafson, 1979). Explanations are usually teleological, i.e. in terms of the agent's intentions. Actions are followed through time via a narrative presentation. The narrative conveys a "story" of what happens. Person, agency, purpose, value, and temporal continuity are key concepts. The agent's description of events is primary data, the object of the scholar being to explain or understand the actions in the agent's own way.

IV.

My thesis then is that the humanistic mode of inquiry is underemployed in evaluation and that the future evaluation of Follow Through could profitably be turned in this direction. Several more specific approaches to evaluation are based upon humanistic reasoning--art criticism, professional review, quasi-legal, and especially case study. (Interestingly, law was a main study of the humanists during the Renaissance.) So such evaluations are being conducted, even though they are not usually consciously tied to the humanities. It is not difficult to see that case studies, for example, have many features in common with history or even literature, the paradigm example of the humanities.

One study of a Follow Through program which exhibits many of the features of humanistic inquiry (though the authors may not agree with that assessment) is a study of the Bank Street Follow Through Project (Zimiles and Mayer, 1980). Bank Street College attempted to implement its model in 54 schools in 14 school districts over an 11 year period of time. In this study Zimiles and Mayer visited 8 of the 14 sites and interviewed local FT directors (12), the staff developers (27), teachers (77), assistant teachers (23), anxillary staff (25), principals (18), parents (15), and Bank Street staff members (10). They also observed 6 classes in 23 schools. The interviews were not structured but were designed to elicit the experiences of these participants about the viability and validity of the Bank Street model and to ascertain its influence on teachers and students.

From these interviews the researchers constructed a story of what happened in these sites. So although the study has some features of a survey, it deliberately eschews percentages and numbers and reports the experiences in the participants' own words. From this emerges a picture of the Bank Street Follow Through project that is far more informative and contains more implications for action than do the equivocal results of the original evaluation.

For example, the great haste with which the Follow Through program was enacted caused a number of far reaching problems. Sites were chosen for the wrong reasons. The decision to participate in Follow Through was made by central administrators in the districts. The principals were excluded, and teachers were sometimes forced to teach in the program against their will. The year to year uncertainties of funding continued

to plague the project. Regular Bank Street staff were reluctant to work on the project. Hence, it was necessary to recruit people from the outside who were less experienced and knowledgeable about the Bank Street approach.

Many "field advisors" were rushed into service unprepared. Their own training was informal and extremely haphazard, which in turn caused them to be defensive when they went to the cooperating sites to train the staff there. The field advisors themselves enacted different roles on site. Some acted as advocates of children, some as curriculum builders, and some as political facilitators. Site variation was great, and results varied dramatically from one site to another. Things seemed to work better in the lower grades and in smaller communities where project people were more in control of events. Gradually, as the field advisors become more experienced, they changed more from training to collaboration with the teachers.

From the beginning everyone had difficulty with the complexity of the Bank Street model. Although the field advisors were deeply committed to the Bank Street approach, they felt the model was too sophisticated for the teachers, and many teachers were too conservative for the model. As Zimiles and Mayer note, the very notion of having a "model" implies discrete practices, a kit. The idea of a visible, picturesque, easily learnable model was not an idea Bank Street subscribed to. The teachers on site, many of whom were improperly chosen by local administrators, had difficulty grasping the Bank Street approach. They saw the Bank Street trainers as too theoretical and too platitudinous. Said one teacher,

"They just philosophized; it wasn't helpful". (Zimiles and Mayer, 1980, p.47.) The teachers much preferred practical "hands-on" workshops.

Under pressure themselves, the trainers often violated Bank Street tenets with the teachers. "There is an arrogance to Bank Street; they think they have all the answers--the 'Kingdom of God is in New York'." (Zimiles and Mayer, p.52.) Some teachers resented the "know-it-all experts". Many teachers mistakenly believed that the trainers had no teaching experience themselves. Obviously, strong emotions were released in the training.

In spite of this, even the most rejecting teachers respected the Bank Street approach: "The children love to come to school, and they're not afraid of criticism" (p.66). Most disturbing to the teachers was loss of control: "The kids go nutty if you open up too much. Kids come in here high as kites" (p.62). Visits to Bank Street College itself made the teachers feel close to the College, and they recognized the special quality of the students and faculty there.

At the same time the teachers wondered about the similarity and needs of their own students. Bank Street staff seemed to ignore regional and social class differences. Students at Bank Street were upper middle-class compared to the ones from impoverished backgrounds that the Follow Through teachers taught. Public schools were more formal institutions with larger class sizes. "The model is more relaxed about reading than we can be in public schools. The children at Bank Street catch up; our children don't catch up. You can't afford to relax" (p.65). Conflicts between Bank Street and the schools over such things as placement of students were the rule. From the Zimiles and Mayer study one has a good picture of what the teachers

were thinking and what they were doing in the classroom.

It is far less clear what the ancillary staff--the social workers, psychologists, etc.--and the parents were doing and thinking. Those parts of the study consist of conclusions without the rich supporting evidence and quotations typical of other sections. The classroom observations, by, contrast, provide extensive examples of what teachers were doing, concluding that overall the intellectual stimulation was not high. Teachers focused on academic skills rather than imaginative enterprises.

The staff developers--those local district personnel responsible for development on site--were chosen by local administrators often without careful consideration of their effectiveness. Local politics prevailed. The staff developers themselves expressed reservations about their own ability to handle the Bank Street model in the classroom. They too complained about the lack of explicitness in the model and the attitude of the Bank Street staff, but they very much liked the idea of "sponsorship," of working with a sponsor like Bank Street. Some noted the lack of commitment and energy on the part of their teachers and wondered whether the model was appropriate for all teachers and children.

According to Zimiles and Mayer, though, the exclusion of the school principals from the implementation was the most critical error. Most principals became passively resistant, often citing regulations as to why things could not be done. This made a dramatic difference on many sites. The principals themselves admired the dedication of the Follow Through teachers but disliked their separateness. Participation in Follow Through seemed to remove the teachers from the principal's control. "The school administration loved to see Bank Street make mistakes," said

one principal (p.130). It was the failures that were remembered.

The section about program effects in the Zimiles-Mayer study seems less satisfying than those detailing the dynamics of the program. When they asked respondents about the distinctive behavior of the students, a nurse found them outgoing, undisciplined, chatty, questioning, independent. A psychologist reported, "They are verbal, curious, problem solve, show initiative, and have a hard time sitting still" (p.150). One mother said, "My kids who were not in Follow Through did not ask questions or talk to a teacher, even when they didn't understand" (p.150). Another psychologist: "The children have a good sense of self and are verbal....They expect respect and accountability, they expect fairness and justice; they view themselves as learners and as important individuals...." (p.151).

Apparently these characteristics did not necessarily endear the students to teachers in later grades. The difficulty of transition to 4th and 5th grades was a major topic of discussion among the teachers. Enthusiasm, impulsivity, individualism, not raising ones hand, conflicted with traditional teachers' notions of how students should behave. (Perhaps this conflict was the best evidence that the Bank Street students were different somehow.) As the authors note, some of these effects might be fused with other causes, such as birth order of the children or observer expectancy. Being verbal, independent, and free is what one might expect from Bank Street students. All in all, one would want these program effects to be more conclusively documented and sorted out.

The authors conclude their study by summarizing the limitations of the Bank Street model. Because the program ideology is not universally accepted, because the teachers must work hard and be of high caliber,

and because the model is costly and sometimes incompatible with bureaucratic demands, it may never be widely implemented. Successful implementation requires certain factors noted in the study, factors not always present in Follow Through. On the other hand, when the model is successfully implemented, the authors conclude, it produces effects in children something like what Bank Street claims for it. The appropriateness of the model for these Follow Through students remains open.

Thus concludes the study of the Bank Street Follow Through project. I have reconstructed the above findings in chronological order. The Zimiles-Mayer study actually presents the information more in terms of categories of respondents, i.e. teachers, principals, etc. It is evident from this brief summary that the study contains a wealth of information and insights not available in the original FT evaluation. Furthermore, the implications for action are far clearer and more intelligible. This is because the study is expressed in terms of action and intentionality, much like the language in which I talked about my children's experience. This is the language of our actions in the real world.

In other words, the study has many of the strengths of humanistic inquiry. It also has some of the weaknesses. The authors were employed by Bank Street, raising the issue of bias. The authors themselves contend that their report is negatively biased against FT since they continually asked for and probed the weaknesses of the programs in their interviews. The fact that so many weaknesses are reported lends the study credibility. If the study had been more positive, it would have been less believable. Subjectivity of reporting is always an issue. (Of course, the original FT evaluation was not without bias problems of its own.) My own experience in educational innovation suggests that other FT sponsors had similar pro-

problems of implementation but these remain unreported.

A related problem with a study like this is how much to rely on the reports of observers about program effects. The study leaves one not totally persuaded here. Were these effects representative of the students as a whole? Were they caused by the program? Unfortunately, the original scientific evaluation does not answer these questions satisfactorily either. Finally, the study is long--209 pages. Nonetheless, in spite of these weaknesses, one gains a wealth of insight from this study simply not available from the original evaluation. For the guidance of future human action, one would choose the humanistic study over the scientific one.

References

Nicola Abbagnano. Humanism in The Encyclopedia of Philosophy. New York: Macmillan, 1967. Vol. 4, p. 69-72.

Isaiah Berlin. Against the Current: Essays in the History of Ideas New York: Viking, 1980.

Norm Chomsky. Language and Responsibility. New York: Pantheon, 1979.

Frederick A. Olafson. The Dialectic of Action. University of Chicago Press, 1979.

Roger Scruton. Humane Education. The American Scholar, Autumn, 1980, pp. 489-498.

Herbert Zimiles and Rochelle Mayer with Elaine Wickens. Bringing Child-Centered Education to the Public Schools. New York: Bank Street College of Education, 1980. 209 pages.